AUSTRALIA

Patent Office
Canberra

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 2210 for a patent by DIATECH PTY LTD filed on 6 March 1998.

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

WITNESS my hand this Seventh day of April 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

AUSTRALIA

Patents Act 1990

DIATECH PTY LTD

PROVISIONAL SPECIFICATION

Invention Title:

V-like Domain Binding Molecules

The invention is described in the following statement:

V-like Domain Binding Molecules

Field of the Invention

The present invention relates to V-like Domain binding molecules with affinities for target molecules. The present invention also relates to compositions comprising these V-li ke domain binding molecules and to methods of diagnosis or treatment which involve the use of these molecules. The present invention also relates to a method for selecting V-like Domain binding molecules with novel binding affinities.

Background of the Invention

Immunoglobulin Superfamily - Ligand Binding

Antibodies are the paradigm of specific high-affinity binding reagents and provide an antigen binding site by interaction of variable heavy (V_H) and variable light (V_L) immunoglobulin domains. The binding interface is formed by six surface polypeptide loops, termed complementarity determining regions (CDRs), three from each variable domain, which are highly variable and combined provide a sufficiently large surface area for interaction with antigen. Specific binding reagents can be formed by association of only the V_H and V_L domains into an Fv module. Bacterial expression is enhanced by joining the V-domains with a linker polypeptide into a single-chain scFv molecule. "Humanisation" of recombinant antibodies by grafting murine CDR loops onto a human Fv framework is disclosed by Winter et al EP-239400. Several attempts to engineer high-affinity single domain binding reagents using either the V_H or V_L domains alone have been unsuccessful, due to lack of binding specificity and the inherent insolubility of single domains in the absence of the hydrophobic face where the V_H and V_L domains interact. V-like Domain binding molecules which have high affinity for target molecules and which are soluble are therefore desirable.

CTLA-4 and CD28

Cytotoxic T-lymphocyte associated antigen 4 (CTLA-4) and a homologous cell-surface protein, CD28, are involved in T-cell regulation during the immune response. CTLA-4 is a 44 kDa homodimer expressed primarily and transiently on the surface of activated T-cells, where it interacts with CD80 and CD86 surface antigens on antigen presenting cells to effect regulation of the immune response (Waterhouse et al.1996, van der Merwe et al. 1997). CD28 is a 44kDa homodimer expressed predominantly

on T-cells and, like CTLA-4, interacts with CD80 and CD86 surface antigens on antigen presenting cells to effect regulation of the immune response (Linsley et al.1990). Current theory suggests that competition between CTLA-4 and CD28 for available ligands controls the level of immune response, for example, gene deletion of CTLA-4 in knock-out mice results in a massive over-proliferation of activated T-cells (Waterhouse et al. 1995).

Each CTLA-4 monomeric subunit consists of an N-terminal extracellular domain, transmembrane region and C-terminal intracellular domain. The extracellular domain comprises an N-terminal V-like domain (VLD; of approximately 14 kDa predicted molecular weight by homology to the immunoglobulin superfamily) and a stalk of about 10 residues connecting the VLD to the transmembrane region. Recent structural and mutational studies on CTLA-4 suggest that binding to CD80 and CD86 occurs via the VLD surface formed from A'GFCC' V-like beta-strands and also from the highly conserved MYPPPY sequence in the CDR3-like surface loop (Peach et al. 1994; Morton et al. 1996; Metzler et al. 1997). Dimerisation between CTLA-4 monomers occurs through a disulphide bond between cysteine residues (Cys^{120}) in the two stalks, which results in tethering of the two extracellular domains, but without any apparent direct association between V-like domains (Metzler et al. 1997). Dimerisation appears to contribute exclusively to increased avidity for the ligands.

In vitro Expression of Soluble Forms of CTLA-4.

Neither the extracellular domains nor V-like domains (VLDs) of human CTLA-4 molecule have been successfully expressed as soluble monomers in bacterial cells, presumably due to aggregation of the expressed proteins (Linsley et al, 1995). Expression of the extracellular N-terminal domain (Met^1 to Asp^{124}, including Cys^{120}) in *E.coli* results in production of a dimeric 28 kDa MW protein, in which two CTLA-4 V-like domains are joined by a disulphide linkage at Cys^{120}. Truncation at Val^{114} removes these cysteines and was intended to enable expression of a 14 kDa VLD in soluble, monomeric form. However, the product aggregated and it was concluded that hydrophobic sites, which were normally masked by glycosylation, were now exposed and caused aggregation (Linsley et al, 1995).

There have been some reports of successful expression of monomeric, glycosylated CTLA-4 extracellular domains in eukaryotic expression systems

(ie CHO cells and the yeast Pichia pastoris; Linsley et al. 1995; Metzler et al. 1997; Gerstmayer et al. 1997). Glycosylation in these eukaryotic expression systems is presumed to occur at the two N-linked glycosylation sites in the VLD (Asn76 and Asn108). However, high yields have only been described for expression of a gene encoding a CTLA-4 VLD fused to Ig-CH2/CH3 domains which produces a dimeric recombinant protein with 2 CTLA-4 VLDs attached to an Fc subunit (WO 95/01994 and AU 16458/95). AU 60590/96 describes mutated forms of CTLA-4 VLDs with single amino acid replacements of the first tyrosine (Y) in the MYPPPY surface loop which retain and modifies the affinity for the natural CD80 and CD86 ligands. AU 60590/96 describes the preferred soluble form of CTLA-4 VLDs as a recombinant CTLA-4/Ig fusion protein expressed in eukaryotic cells and does not solve the aggregation problem in prokaryote expression systems. EP 0757099A2 describes the use of CTLA-4 mutant molecules, for example the effect of changes on ligand binding of mutations in the CDR3 loop.

Summary of the Invention

The present inventors have now developed novel binding molecules derived from the V-like domains (VLDs) of immunoglobulin superfamily members such as CTLA-4 and CD28. Replacement of loop structures within the VLDs results unexpectedly in the production of monomeric, correctly folded molecules with altered binding specificities and improved solubility.

Accordingly, in a first aspect the present invention provides a V-like domain (VLD) from a member of the immunoglobulin superfamily, in which at least one surface loop structure or part thereof is modified or replaced such that the binding affinity of the VLD is altered and the solubility of the modified VLD is improved when compared with the unmodified VLD.

The phrase "V-like domain" or "VLD" is intended to refer to a domain which has similar structural features to the variable heavy (V_H) or variable light (V_L) antibody domains. The similar structural features of these immunoglobulin V-like domains include CDR loop structures. By "CDR loop structures" we mean surface polypeptide loops like the complementarity determining regions in antibody V-domains.

The phrase "immunoglobulin superfamily" is intended to refer to any molecule which has a "V-like domain". Examples of members of the immunoglobulin superfamily are T-cell surface proteins such as CTLA-4,

CD28. It will be appreciated by those skilled in the art that other members of the immunoglubulin superfamily which can provide homologous V-like domains suitable for the invention are other T-cell surface proteins CD2, CD4, CD7 and CD16; B cell surface proteins including CD19, CD79a, CD22, CD33, CD80 and CD86; adhesion molecules such as CD48, CD541CAM and CD58. These molecules, which are listed in Table 1, provide a non-exhaustive list of structures which may form the basis for the single domain binding molecules of the present invention.

TABLE 1: IMMUNOGLOBULIN SUPERFAMILY MEMBERS

Molecule	Size	Structure
Tcell lymphocyte receptors		
CD2	45-58kDa	VC[1] domains
CD4	55kDa	V2C2
CD7	40kDa	V domain
CD16	50-65kDa	2x C domains
B cells		
CD19	95kDa	2x C domains
CD79a	33kDa	
CD22	130-140kDa	1xV 6xC domains
CD33	67kDa	VC domain
CD80	60kDa	VC domain
CD86	60kDa	VC domain
Adhesion molecules		
CD48	45kDa	VC domain
CD54ICAM	85-110kDa	
CD58	55-70kDa	VC domain

1 V = variable Ig domain, C = constant domain

These molecules are discussed in (1) The Leucocyte Antigen Facts Book, 1993, Eds Barclay et al., Academic Press, London; and (2) CD Antigens 1996 (1997) Immunology Today 18, 100-101, the entire contents of which are incorporated herein by reference.

It will be appreciated that the VLDs of the present invention may be coupled together, either chemically or genetically, to form multivalent or multifunctional reagents. For example, the addition of c-terminal tails, such as in the native CTLA-4 with Cys^{120}, will result in a dimer.

The VLDs of the present invention may also be coupled to other molecules for various diagnostic formulations. For example, the VLDs may include a c-terminal polypeptide tail or may be coupled to streptavidin or biotin for multi-site *in vitro* assays. The VLDs may also be coupled to radioisotopes, dye markers or other imaging reagents for *in vivo* detection and/or localisation of cancers, blood clots, etc. The VLDs may also be immobilised by coupling onto insoluble devices and platforms for diagnostic and biosensor applications.

In a preferred embodiment of the first aspect of the present invention, the V-like domain is derived from a molecule other than a variable heavy (V_H) or variable light (V_L) antibody domain or corresponding V-alpha or V-beta T-cell receptor domain. Preferably, the V-like domain is derived from a T-cell surface protein such as CTLA-4 or CD28.

In a most preferred embodiment of the first aspect of the present invention, the V-like domain is derived from the extracellular domain of the CTLA-4 molecule or the CD28 molecule. In a further preferred embodiment one or more surface loops of the CTLA-4 V-like domain and preferably the CDR-1, CDR-2 or CDR-3 loops are replaced with a polypeptide which has a binding affinity for a target molecule of interest. Target molecules of interest include, but are not limited to, drugs, steroids, pesticides, antigens, growth factors, tumour markers, cell surface proteins or viral coat proteins. It will be appreciated that these VLDs will be polyspecific, having affinities directed by both their natural surfaces and modified polypeptide loops.

In a further preferred embodiment the effect of replacing or modifying the CTLA-4 or CD28 V-like domain surface loops is to abolish the natural affinity to CD80 and CD86.

In a further preferred embodiment, one or more of the CDR loops is replaced with a peptide hormone, such as somatostatin which is a 14 residue intra-disulphide bonded polypeptide important in cancer cell recognition, or with a viral protein such as the human influenza virus haemagglutinin protein.

In a further preferred embodiment the V-like domain of the binding molecule includes CDR loop structures homologous in character to CDR loops found in camelid antibodies. For example, the CDR loops may contain one or more non-conventional substitutions (eg. hydrophobic to polar in nature). In another preferred embodiment, the CDR-1 and CDR-3 loops may adopt non-canonical conformations which are extremely heterologous in length. The V-like domain may also possess a disulphide linkage interconnecting the CDR-1 and CDR-3 loops.

In a second aspect the present invention provides a polynucleotide encoding a VLD binding molecule of the first aspect of the present invention. The polynucleotide may be incorporated into a plasmid or expression vector.

In a third aspect the present invention provides a prokaryotic or eukaryotic host cell transformed with a polynucleotide according to the second aspect of the present invention.

In a fourth aspect the present invention provides a method of producing a VLD molecule which includes culturing a host cell according to the third aspect of the present invention under conditions enabling expression of the VLD molecule and optionally recovering the VLD molecule.

In a preferred embodiment of the present invention the VLD molecule is produced by expression in a bacterial host. Preferably, the VLD molecule is unglycosylated.

In a fifth aspect the present invention provides a pharmaceutical composition including a VLDs molecule of the first aspect of the present invention and a pharmaceutically acceptable carrier or diluent.

In a sixth aspect the present invention provides a method of treating a pathological condition in a subject, which method includes administering to the subject a VLDs molecule according to the first aspect of the present invention.

For *in vivo* applications it is preferable that VLDs molecules are homologous to the subject of treatment or diagnosis and that any possible xenoantigens are removed. Accordingly it is preferred that VLD molecules for use in clinical applications are substantially homologous to naturally occurring human immunoglobulin superfamily members.

In a seventh aspect the present invention provides a method of selecting a VLD molecule with a binding affinity for a target molecule which includes

(i) producing a library of polynucleotides encoding V-like domains from one or more members of the immunoglobulin superfamily, wherein the V-like domains include random substitutions in the CDR loop structures; and

(ii) screening the library for expression of a V-like domain with a binding affinity for the target molecule.

In a preferred embodiment of the seventh aspect of the present invention, the screening process involves displaying the substituted V-like domains as a gene III protein fusions on the surface of bacteriophage particles. The library may comprise bacteriophage vectors such as pHFA or fd-tet-dog containing the polynucleotides encoding the V-like domains.

Brief description of the Figures

Figure 1. Somatostatin peptide.

Figure 2. Designations and amino acid sequences of recombinant CTLA-4 STMs.

Figure 3. Gel filtration profile of recombinant CTLA-4 STMs.

Figure 4. Expression characteristics of CTLA-4 STM proteins.

Figure 5. Gel filtration profiles of recombinant CTLA-4 STM proteins.

Figure 6. CTLA-4 secondary structure and examples of random libraries of STMs.

Detailed Description of the Invention

The present invention relates to the design of novel soluble VLDs binding molecules derived from the V-like domain of immunoglobulin superfamily members, such as the human CTLA-4 molecule. The preferred binding molecules of the present invention provide the following advantages (i) use of a native human protein obviates the need for subsequent humanisation of the recombinant molecule, a step often required to protect against immune system response if used in human treatment; (ii) the domain is naturally monomeric as described above, yet simple incorporation of residue Cys^{120} results in production of a tethered dimeric molecule; and (iii) structural modifications have resulted in improved bacterial expression levels.

Prior to publication of the first CTLA-4 structure determination, available sequence data and mutational analyses of both this molecule and CD28 were analysed. This allowed modelling and prediction of the regions

corresponding to CDR1, 2 and 3. It was hypothesised that such areas would be susceptible to mutation or substitution without substantial effect upon the molecular framework and hence would allow expression of a correctly folded molecule. The subsequently published structure (Metzler et al. 1997) showed these predictions to be accurate, despite the unexpected separation of CDR1 from the ligand-binding site, and the extensive bending of CDR3 to form a planar surface contiguous with the ligand binding face.

In an initial set of experiments the V-like domain of the human CTLA-4 molecule was modified by replacement of CDR loops with either of two defined peptides. The two peptides were somatostatin and a portion of the human influenza virus haemagglutinin protein (HA-tag). Somatostatin (SRIF: somatotropin release-inhibiting factor) is a 14 residue peptide comprising a disulphide bond that forces the central 10 residues into a loop (Figure 1). Somatostatin is biologically widespread within the body and mediates a number of diverse physiological functions such as regulation of growth hormone secretion etc (Reisne, 1995). Somatostatin binds a number of specific receptors (there are at least five subtypes) which have differing tissue specificities and affinities (Schonbrunn et al. 1997). These aspects of binding and activation are as yet poorly understood, but one salient feature is the high density of somatostatin receptors present on a number of cancerous cell lines, for example cancers of the neuro-endocrine system and small lung cancers (Reubi 1997). Artificial analogues of somatostatin have been produced for imaging of such tumours which are resistant to degradation compared with the highly labile somatostatin peptide.

The haemagglutinin epitope sequence consists of the 9 residues YPYDVPDYA. A commercially produced antibody is available which specifically recognises this sequence. The epitope tag can be detected when randomly or directionally incorporated within the structure of proteins (Canfield et al. 1996).

Replacement of one or more CDR loops in the CTLA-4 V-like domain with somatostatin or the HA-tag resulted in the production of soluble, monomeric, unglycosylated binding molecules. This surprising finding shows that V-like domains provide a basic framework for constructing soluble, single domain molecules, where the binding specificity of the molecule may be engineered by modification of the CDR loop structures.

The basic framework residues of the V-like domain may be modified in accordance with structural features present in camelid antibodies. The camel heavy chain immunoglobulins differ from "conventional" antibody structures by consisting of only a single VH domain (Hamers-Casterman et al. 1993). Several unique features allow these antibodies to overcome the dual problems of solubility and inability to present a sufficiently large antigen binding surface.

First, several non-conventional substitutions (predominantly hydrophobic to polar in nature) at exposed framework residues reduce the hydrophobic surface, while maintaining the internal beta-sheet framework structure (Desmyter et al. 1996). Further, within the three CDR loops several structural features compensate for the loss of antigen binding-surface usually provided by the V_L domain. While the CDR2 loop does not differ extensively from other V_H domains, the CDR-1 and -3 loops adopt non-canonical conformations which are extremely heterologous in length. For example, the H1 loop may contain anywhere between 2-8 residues compared to the usual five in Ig molecules. However, it is the CDR3 loop which exhibits greatest variation: in 17 camel antibody sequences reported, the length of this region varies between 7 and 21 residues (Muyldermans et al. 1996). Thirdly, many camelid VH domains possess a disulphide linkage interconnecting CDRs-1 and -3. The function of this structural feature appears to be maintenance of loop stability and providing a more contoured, as distinct from planar, loop conformation which both allows binding to pockets within the antigen and gives an increased surface area. However, not all camelid antibodies possess this disulphide bond: the one published example of a llama antibody lacks this feature, suggesting that it is not an absolute structural requirement (Spinelli et al 1996).

These foregoing features have enabled camelid V-domains to present as soluble molecules in vivo and with sufficiently high affinity to form an effective immune response against a wide variety of target antigens. In contrast, cell surface receptors of the Ig superfamily (such as CD4 and CD2) comprise V-like binding domain and appear to bind cognate receptors with surface features other than the CDR loops. These V-like domains bind to cognate receptors with very low affinity. Physiological signalling between 2 cells are mediated by avidity of multi-point binding, when 2 cell surfaces connect and each contains multiple receptors. CD2 is a well-characterised

example: binding to CD58 is mediated by a highly constrained set of minor electrostatic interactions generated by charged and polar residues located in the GFCC'C" face (not the CDR-type loops). This results in a low affinity but highly specific interaction (Bodian et al 1994).

The present invention also relates to a method for generating and selecting single VLD molecules with novel binding affinities for target molecules. This method involves the application of well known molecular evolution techniques to V-like domains derived from members of the immunoglobulin superfamily. The method may involve the production of phage display libraries for screening large numbers of mutated V-like domains.

Filamentous Fd-bacteriophage genomes are engineered such that the phage display, on their surface, proteins such as the Ig-like proteins (scFv, Fabs) which are encoded by the DNA that is contained within the phage (Smith, 1985; Huse et al., 1989; McCafferty et al., 1990; Hoogenboom et al., 1991). Protein molecules can be displayed on the surface of Fd bacteriophage, covalently coupled to phage coat proteins encoded by gene III, or less commonly gene VIII. Insertion of antibody genes into the gene III coat protein gives expression of 3-5 recombinant protein molecules per phage, situated at the ends. In contrast, insertion of antibody genes into gene VIII has the potential to display about 2000 copies of the recombinant protein per phage particle, however this is a multivalent system which could mask the affinity of a single displayed protein. Fd phagemid vectors are also used, since they can be easily switched from the display of functional Ig-like fragments on the surface of Fd-bacteriophage to secreting soluble Ig-like fragments in E. coli. Phage-displayed recombinant protein fusions with the N-terminus of the gene III coat protein are made possible by an amber codon strategically positioned between the two protein genes. In amber suppressor strains of *E. coli*, the resulting Ig domain-gene III fusions become anchored in the phage coat.

A selection process based on protein affinity can be applied to any high-affinity binding reagents such as antibodies, antigens, receptors and ligands (see, for example, Winter and Milstein, 1991, the entire contents of which are incorporated herein by reference). Thus the selection of the highest affinity binding protein displayed on bacteriophage is coupled to the recovery of the gene encoding that protein. Ig-displaying phage can be

affinity selected by binding to cognate binding partners covalently coupled to beads or adsorbed to plastic surfaces in a manner similar to ELISA or solid phase radioimmunoassays. While almost any plastic surface will adsorb protein antigens, some commercial products are especially formulated for this purpose, such as Nunc Immunotubes.

Although there are several alternative approaches to modify binding molecules the general approach for all phage displayed proteins conforms to a pattern in which individual binding reagents are selected from phage display libraries by affinity to their cognate receptor. The genes encoding these reagents are modified by any one or combination of a number of in vivo and in vitro mutation strategies and constructed as a new gene pool for phage display and selection of the highest affinity binding molecules.

In order that the nature of the present invention may be more clearly understood preferred forms thereof will now be described with reference to the following examples.

Examples

1. Experimental Methods

CTLA-4 STM (STM: soluble truncated mutants of CTLA-4, used herein to describe CTLA-4 V-like domains which may include a c-terminal polypeptide tail) gene construction and cloning was by standard and well-described techniques (Polymerase chain reaction with specifically designed oligonucleotide primers, splice overlap extension, restriction enzyme digests etc). A list of oligoucleotide primers used is given in Table 2. The wild-type construct was amplified from cloned human CTLA-4 DNA (and could be similarly amplified from reverse transcribed human cDNA by a competent worker in the field) using the oligonucleotide primers #3553 and #4316, which incorporated SfiI and NotI restriction sites at the 5' and 3' ends respectively. These terminal primers were used in all further constructions except where primer #4851 was used to incorporate an ApaL1 site at the 5' end, and where the 3' end was cloned to include a c-terminal tail including residue Cys^{120} using oligonucleotide #4486. For somatostatin and haemagglutinin loop replacements, splice overlap PCR reactions were performed using primers #4585, #4766, and #4586, followed by final extensions as for the wild type STM. This strategy and combinations of these oligonucleotide primers were used to produce all CDR loop replacement variations.

TABLE 2
TABLE OF OLIGONUCLEOTIDES

OLIGONUCLEOTIDE	NUMBER	SIZE	SEQUENCE[1]	FEATURES
For 5' CTLA-4 amplification	3553	54	TTATTACTCgCggCCCAgCCggCCATggCCgCAATgCACgTggCCCAgCCTgCT	forward primer + SfiI site
	4851	39	TCTCACAgTgCACAggCAATgCACgTggCCCAgCCTgCT	forward primer + ApaL1 site
For 3' CTLA-4 amplification	4316	42	ATCTgCggCCgCTACATAAATCTggaTACCgTTgCCgATgCC	reverse primer + NotI site
	4486	66	gCTgAATTCTgATCAgTgATggTgATggTgATgTgCggCCgCgTCAgAATCTggCACggTTCTgg	reverse primer for Cys120 incorporation
For CDR1-somatostatin	4585	84	AgCTTTgTgTgTgAgTATgCAgCTggCTgCAAgAATTTCTTCTggAAgACTTTCACATCCTgTgCCACTgAggTCCg ggTgACA	forward primer + Som
For CDR-1 randomisation	4483	66	AgCTTTgTgTgTgAgTATgCAN[2]Ng/TNNg/TNNg/TNNg/TNNg/TNNg/TNNg/TNNg/TgCCACTgAggTCCggTgA CA	forward primer + 8 residue random loop
	4484	68	CACgTggCCCAgCCTgCTgTggTACTggCCAgCAgCCgAggCATCgCCAgCTTTgTgTgTgAgTATgC	extension primer
	4254	66	gTgTgTgAgTACgCgTNCNNg/CNNg/CNNg/CNNg/CNNg/CNNg/CTgCNNg/CgCTACTgAggTTCgTgTgACCgTC	forward primer with MluI and AflII sites, extended randomised loop with conserved cysteine
For CDR1-som-randomisation	4835	75	AgCTTTgTgTgTgAgTATgCAgCTggCTgCAAgAATNNg/TNNg/TNNg/TNNg/TNNg/T NNg/TACATCCTgTgCCACTgAggTC	
For CDR2 haemagglutinin tag	4766	21	gTAggTTgCCgCACAgACTTC	back primer for splice overlap
	4775	66	gAAgTCTgTgCggCAACCTACCCgTATgACgTTCCgACTACgCCCTAgATgATTCCATCTgCACg	forward primer: tag + overlap
For CDR2 randomisation				
For CDR3-somatostatin	4586	84	CTgggTACCgTTgCCgATgCCACAggATgTgAAAgTCTTCCAgAAgAAATTCTTgCAgCCAgCCTCCACCTTgCAg ATgTAgAg	reverse primer +Som
For CDR3 randomisation	4482	69	CTgggTACCgTTgCCgATgCCC/ANNC/ANNC/ANNC/ANNC/ANNC/ANNC/ANNC/ANNC/ANNGTCCACCTTg CAgATgTAgAg	reverse primer + 9 residue random loop
	4275	67	AggTggAANNg/CNNg/CNNg/CNNg/CNNg/CNNg/CTgCNNg/CNNg/CNNg/CNNg/CNNg/CNNg/CNNg/CggCAT CggCAACggTAC	forward primer withPstI and KpnI overhangs, long random loop replacement incorporating conserved cysteine for bridge
For CDR3-som-randomisation	4836	75	CTgggTACCgTTgCCgATgCCACAggATgTA/CNNA/CNNA/CNNA/CNNA/CNNA/CNNATTCTTgCAgCCAgCC TCCACCTTgCA	

[1] All oligonucleotides are described 5' to 3'.

2 N represents combination of the four nucleotides

For randomisation of the six residues at the tip of the somatostatin molecules which replaced CDR loops 1 and 3, similar splice-overlap techniques were used, for example primers #4835 and #4836. Randomisation was by synthesis of oligonucleotides with a given triplets being encoded by the sequence NNg/T where N represents any of the four possible bases. This combination covers all possible residues at a given position. A similar approach was used for the replacement of CDR loops directly, by randomised peptides of sizes equalling or greater than the size of the wild type CTLA-4 STM loops. Again, this involved production of suitably randomised oligonucleotides that could be extended by splice-overlap PCR. Examples of such primers are #4483 and #4482. In some instances, a variant technique was used for STM gene construction, where randomised oligonucleotide primers were designed which incorporated restriction sites for direct cloning into the similarly modified (with complementary restriction sites) CTLA-4 STM framework. Such strategies were predominantly used in examples where conserved cysteine residues were incorporated within the random loops, to allow disulphide linkage formation between the newly constructed and extended loop regions. The addition of such added structural feature was to mimic the stability of camelid antibodies which utilise this strategy to give extra rigidity to the extended loop regions. Examples of two such oligonucleotide primers are #4254 and #4275.

Completed constructs (which are described in Figure 2) were cut with the restriction enzymes *Sfi*1 and *Not*1 and cloned into like sites in the E.coli expression vector pGC for production of soluble protein. This vector allows high level expression of heterologous proteins, which are targeted to the periplasmic space by means of a leader sequence. The leader sequence is subsequently cleaved to produce the mature protein. In addition, this vector contains two in-frame 8 residue tag sequences (FLAG tags) which allow affinity purification of the recombinant protein. All DNA constructs were verified by restriction analysis and DNA sequencing and tested for expression of recombinant protein by standard and well-understood techniques (Polyacrylamide gel electrophoresis, Western blot etc). For bacteriophage and phagemid display, completed STM constructs were cut with the restriction enzymes *Sfi*I and *Not*I or *Apa*LI and *Not*I and cloned into the

vectors pHFA or pFD-Tet-DOG respectively. These vectors allow display of the STMs on the surface of bacteriophage.

Recombinant proteins were purified by the following method, which is but one variation of well established techniques. Bacterial clones were grown overnight in 2YT medium/37°C /200 rpm/100mg/ml ampicillin, 1% glucose (final). Bacteria were diluted 1/100 into either 0.5 or 2l of 2YT medium supplemented with 100mg/ml ampicillin, 0.1% glucose (final), and grown at 28°C/ 200 rpm. These cultures were grown to an optical density of between 0.2-0.4, at which stage they were induced with 1mM IPTG (final). Cultures were allowed to grow for 16 hours (overnight) before harvesting.

Bacteria were collected by centrifugation (Beckman JA-14 rotor or equivalent/6K/10min/4°C) and the periplasmic fraction collected by standard techniques. Briefly, this involved resuspension of bacterial pellets in a 1/25th volume of spheroplast forming buffer consisting of 100mM Tris-HCl/0.5M sucrose/0.5 mM EDTA (pH8.0), followed by addition of 1/500th volume of hen egg lysozyme (2mg/ml in water) and incubation for 10min. A 0.5x solution of the above spheroplasting buffer was then added to a final volume of 1/5th of the original culture, and the incubation continued for a further 20min. The cell debris was then pelleted by centrifugation (Beckman JA-14 rotor or equivalent/9K/15min/4°C) and the supernatant containing the periplasmic fraction collected. All of the above procedures were performed at 4°C. Samples were processed immediately, or if storage was required were stored at 4°C in the presence of protease inhibitors. If freezing was required, no more than one freeze-thaw cycle was allowed.

In preparation for affinity chromatography, samples were subjected to sonication and filtration by standard techniques. Chromatography was by affinity column utilising the anti-flag M2 antibody (Kodak). The eluted fraction was dialysed against 1x TBS (pH8.0)/2x changes, concentrated, and loaded onto a Superose 12 column (Pharmacia Biotech). Fractions corresponding to monomeric, dimeric and tetrameric species were collected, as determined by pre-calibrated elution times, followed by concentration, analysis and storage at -20°C. All of the above protein chemistry methods are standard techniques within the field. Purified proteins were analysed by standard techniques for example polyacrylamide gel electrophoresis, western blot, dot blot.

Cloning and expression in the bacteriophage expression vectors pHFA and fd-tet dog, and subsequent production of recombinant bacteriophage, were by standard and well-established techniques. Screening of libraries of randomised CTLA-4 STMs was by standard and well-established techniques (Galanis et al 1997).

2. Results

An STM of the human CTLA-4 protein containing wild type CDR loops can be expressed in and purified from the *E.coli* periplasmic space. This STM (Figure2) lacks the Cys^{120} that tethers two full-length molecules together into a dimer. Despite the absence of such covalent interactions, superose 12 profiles revealed a preponderance of dimeric and tetrameric compared to monomeric species. A high proportion of aggregated material was also present. A typical superose 12 profile is shown in Figure 3. One possible explanation for the observed oligomerisations is the absence of glycosylation from the bacterially produced protein. *E.coli* is unable to glycosylate the Asn78 residue which is glycosylated in eukaryotic expression systems, and possibly plays a role in masking hydrophobic patches on the CTLA-4 surface.

In an initial set of experiments either the CDR1 or the CDR3 loop of this CTLA-4 STM was replaced with the somatostatin peptide (Figures 2, 4). This 14 residue peptide is conformationally constrained by an intra-disulphide linkage between Cys^{3} and Cys^{14} (Figure 1). We reasoned that this would form a discrete protein loop, analogous to the CDR loops found in antibodies, and particularly analogous to the long CDRs camelid antibodies which are also stabilised by a disulphide linkage. We also compared the effect of substituting CDR1 in the presence or absence of $\mathbf{Cys^{120}}$ ie whether a tethered dimer could be produced. These initial experiments produced an unexpected and surprising result. Substitution of either CDR1 or -3 with somatostatin significantly enhanced the production of monomeric protein. This is illustrated in drawings 3 and 5, where replacement of CDR3 can be seen to significantly increase the ratio of monomeric to dimeric/tetrameric protein species.

Testing of isolated monomeric STM proteins showed that they remained monomeric after several rounds of freeze-thawing. Evidence for retention of the correct CTLA-4 conformation came from blotting studies with an conformationally-specific anti-CTLA-4 antibody. This antibody

blocks binding to the CTLA-4 co-receptors and recognises protein by dot but not by western blot (where the protein is denatured). Monomeric species from both CDR1 and -3 replacements reacted strongly with this antibody. Interestingly, this antibody recognised the wild type monomer and the tethered dimer (CDR1 replaced) only poorly, in contrast to the strong reaction with the modified protein species. This suggests that in the wild type STM some form of local interaction is occurring that occludes the antibody binding site, and that this interaction is similar to the result when two CTLA-4 molecules are tethered together (presumably blocking access to the antibody).

In further experiments, simultaneous replacement of both CDR1 and -3 by somatostatin resulted in high-level production of monomeric protein (Figure 5). This result suggests that extensive alternative loop residues can be accommodated by the CTLA-4 scaffolding. Alternatively, one of the somatostatin loops may lie flat against the face of the molecule in a manner analogous to that of the CDR3 loop of CTLA-4.

In a further extension of the CDR-replacement strategy, a region corresponding to CDR2 was replaced with the 8-residue haemagglutinin (HA) tag sequence. Use of the conformationally constrained somatostatin loop in this position was considered unsuitable as this region partially encompasses the C" beta strand running the length of the molecule. The HA tag could be detected upon this CTLA-4 STM, however, gel filtration revealed a range of protein species, from monomeric through to aggregated species (Figure 5).

Final proof of the principle that the CTLA-4 CDR loops could be replaced with other peptides to produce monomeric, soluble, chaemeric proteins was by simultaneous replacement of all three CDRs with two somatostatin and one HA epitope respectively. This construct produced a correctly folded and monomeric protein upon gel filtration chromatography (Figure 5).

3. Bacteriophage display of CTLA-4 and variants.

CTLA-4 STMs were displayed as fusion proteins with the geneIII protein on the surface of fd-bacteriophage. Two different forms of expression were employed: (i) utilising the phagemid vector pHFA where recombinant bacteriophage are rescued from transformed cells by the addition of helper bacteriophage. The salient feature of this system is that each bacteriophage

carries between 0-1 copies of the STM as geneIII fusion proteins, the remainder (3-5) copies of the geneIII protein are wild type; (ii) to increase the copy number of the STM- geneIII fusion, CTLA-4 STM constructs were engineered into the phage vector fd-tet-dog. This vector encodes the bacteriophage in it's entirety, thus no helper bacteriophage are required and each bacteriophage carries only the recombinant geneIII fusion protein (3-5 copies per bacteriophage). Dot blot analysis using anti-CTLA-4 antibody indicated that correctly folded and comformationally correct CTLA-4 was expressed on the surface of purified bacteriophages.

4. Construction of CTLA-4 randomised libraries

This section describes production of randomised CTLA-4 STMs, where areas of various CDR loops were randomised to produce libraries of novel binding proteins. This follows the results detailed above, which indicate that CDR replacements do not perturb the underlying framework structure, and that the recombinant proteins can be displayed on the surface of bacteriophage, allowing selection/amplification by established library-screening techniques.

Four different examples of randomisation of CDR loops of CTLA-4 STMs were produced, these are listed in Figure 6. These are:

i) Exact replacement of the residues of the CTLA-4 STM CDR1 loop with randomised four residues.

ii) As for **i)** except replacement with a nine residue loop containing a conserved cysteine at position 8.

iii) Simultaneous replacement of CDR1 and -3 loops of CTLA-4 STM with eight residue (CDR1) and nine residue (CDR3) randomised loops.

iv) The CTLA-4 framework proved capable of accommodating dual disulphide-bonded somatostatin peptides. The native somatostatin peptide assumes a conformation where four residues at the tip (Phe-Trp-Lys-Thr) contact the somatostatin receptor. It is therefore likely that these four residues will be accessible in the CTLA-4-somatostatin STM. Thus randomised libraries of STMs were produced by randomising the six residues at the tip of the somatostatin peptide, and incorporating these now randomised peptides into the CDR1 and -3 loops of CTLA-4-Som STM VV3 (Figure 6). Libraries were cloned into both pHFA and fd-tet-dog for

screening. The theoretical number of possible combinations of this library is 2012.

STMs where CDR2 is replaced with the HA peptide in combination with CDR1 and -3 replacements do not compromise folding and solubility characteristics. Thus, any binding molecules identified by screening procedures can be further affinity enhanced by randomisation and selection of the CDR2 loop.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

Dated this sixth day of March 1998

DIATECH PTY LTD
Patent Attorneys for the Applicant:

F B RICE & CO

References

Bodian DL, Jones EY, Harlos K, Stuart DI, Davis SJ (1994) Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 A resolution. Structure 2: 755-766

Canfield VA, Norbeck L, Levenson R (1996) Localization of cytoplasmic and extracellular domains of Na,K-ATPase by epitope tag insertion Biochem. 35: 14165-14172

Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme Nat Struct Biol 3:803-811

Galanis M, Irving RA, and Hudson PJ (1997) Current Protocols in Immunology, 17.1.1-17.1.45.

Gerstmayer B, Pessara U, Wels W (1997) Construction and expression in the yeast Pichia pastoris of functionally active soluble forms of the human costimulatory molecules B7-1 and B7-2 and the B7 counter-receptor CTLA-4. FEBS Lett 407: 63-68

Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R (1993) Naturally occurring antibodies devoid of light chains. Nature 363: 446-448

Holliger, P., Prospero, T., and Winter, G. 1993. "Diabodies": Small bivalent and bispecific antibody fragments. Proc. Natl. Acad. Sci. USA 90: 6444-6448.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and Winter, G. 1991. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. 19:4133-4137.

Huse, W. D.,Sastry, L.,Iverson, S. A.,Kang, A. S.,Alting, M. M.,Burton, D. R.,Benkovic, S. J., & Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda . Science, 246: 1275-81.

Linsley, PS, Clark EA, and Ledbetter JA (1990) The T cell antigen CD28 mediates adhesion with B cells by interacting with the activation antigen, B7/BB1 Proc. Natl. Acad. Sci USA 87: 5031.

Linsley PS, Nadler SG, Bajorath J, Peach R, Leung HT, Rogers J, Bradshaw J, Stebbins M, Leytze G, Brady W, et al (1995) Binding stoichiometry of the cytotoxic T lymphocyte-associated molecule-4 (CTLA-4). A disulfide-linked homodimer binds two CD86 molecules. J Biol Chem 270: 15417-24

McCafferty, J.,Griffiths, A. D.,Winter, G., & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. Nature, 348: 552-4.

Metzler WJ, Bajorath J, Fenderson W, Shaw SY, Constantine KL, Naemura J, Leytze G, Peach RJ, Lavoie TB, Mueller L, Linsley PS (1997) Solution structure of human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28 Nat Struct Biol 4: 527-531

Morton PA, Fu XT, Stewart JA, Giacoletto KS, White SL, Leysath CE, Evans RJ, Shieh JJ, Karr RW (1996) Differential effects of CTLA-4 substitutions on the binding of human CD80 (B7-1) and CD86 (B7-2) J Immunol 156: 1047-1054

Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA, Hamers R (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng 7: 1129-1135

Peach RJ, Bajorath J, Brady W, Leytze G, Greene J, Naemura J, Linsley PS (1994) Complementarity determining region 1 (CDR1)- and CDR3-analogous regions in CTLA-4 and CD28 determine the binding to B7-1 J Exp Med 180: 2049-2058

Reisine, T. (1995) Somatostatin. Cell Molec. Neurobiol. 15: 597-614

Reubi, J. C. (1997) Regulatory peptice receptors as molecular targets for cancer diagnosis and therapy. Q. J. Nucl. Med. 41: 63-70

Schonbrunn A, Gu YZ, Brown PJ, Loose-Mitchell D (1995) Function and regulation of somatostatin receptor subtypes. Ciba Found. Symp. 190: 204-217

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, 228, 1315-1317.

Spinelli S, Frenken L, Bourgeois D, de Ron L, Bos W, Verrips T, Anguille C, Cambillau C, Tegoni M (1996) The crystal structure of a llama heavy chain variable domain Nat Struct Biol 3: 752-757

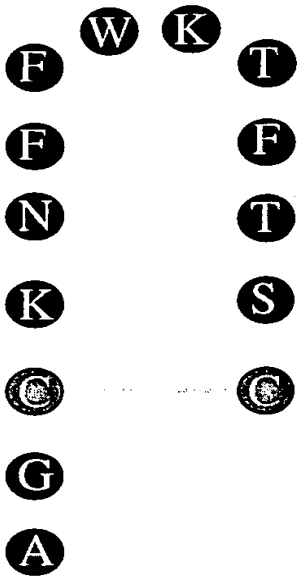
van der Merwe PA, Bodian DL, Daenke S, Linsley P, Davis SJ (1997) CD80 (B7-1) binds both CD28 and CTLA-4 with a low affinity and very fast kinetics. J Exp Med 185: 393-403

Waterhouse P, Marengere LE, Mittrücker HW, Mak TW (1996) CTLA-4, a negative regulator of T-lymphocyte activation. Immunol Rev 153: 183-207

Waterhouse P, Penninger JM, Timms E, Wakeham A, Shahinian A, Lee KP, Thompson CB, Griesser H, Mak TW (1995) Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4 Science 270: 985-988

Winter G, Milstein C. (1991). Man-made antibodies. Nature. 349: 293-299

Figure 1



Somatostatin (somatotropin release-inhibiting factor SRIF) is a cyclic 14-amino acid peptide. The cyclic nature is provided by a disulphide linkage between the cysteine residues at positions 3 and 14. The four residue forming the tip of the loop (Phe-Trp-Lys-Thr) are implicated in binding to members of the somatostatin receptor family.

Figure 2

A

	CDR1	CDR2	CDR3
S2	AMHVAQPAVVLASSRGIASFVCEYA..**SPGK**..ATEVRVTVLRQADSQVTEVCAAT.....	**YMMGNELTF**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**LMYPPPYYL**.....GIGNGAQIYV +21(FLAG)
PP2	AMHVAQPAVVLASSRGIASFVCEYA..**AGCKNFFWKTFTSC**..ATEVRVTVLRQADSQVTEVCAAT.	**YMMGNELTF**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**LMYPPPYYL**.....GIGNGAQIYV +21(FLAG)
PP5	AMHVAQPAVVLASSRGIASFVCEYA..**AGCKNFFWKTFTSC**..ATEVRVTVLRQADSQVTEVCAAT.	**YMMGNELTF**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**LMYPPPYYL**.....GIGNGAQIYVIDPEPCPDSD +21(FLAG)
PP8	AMHVAQPAVVLASSRGIASFVCEYA..**SPGK**..ATEVRVTVLRQADSQVTEVCAAT.....	**YMMGNELTF**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**AGCKNFFWKTFTSC**..GIGNGAQIYV +21(FLAG)
VV3	AMHVAQPAVVLASSRGIASFVCEYA..**AGCKNFFWKTFTSC**..ATEVRVTVLRQADSQVTEVCAAT.	**YMMGNELTF**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**AGCKNFFWKTFTSC**..GIGNGAQIYV +21(FLAG)
XX4	AMHVAQPAVVLASSRGIASFVCEYA..**SPGK**..ATEVRVTVLRQADSQVTEVCAAT.....	**YPYDVPDYA**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**LMYPPPYYL**.....GIGNGAQIYV +21(FLAG)
ZZ3	AMHVAQPAVVLASSRGIASFVCEYA..**AGCKNFFWKTFTSC**..ATEVRVTVLRQADSQVTEVCAAT.	**YPYDVPDYA**..LDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVE..	**AGCKNFFWKTFTSC**..GIGNGAQIYV +21(FLAG)

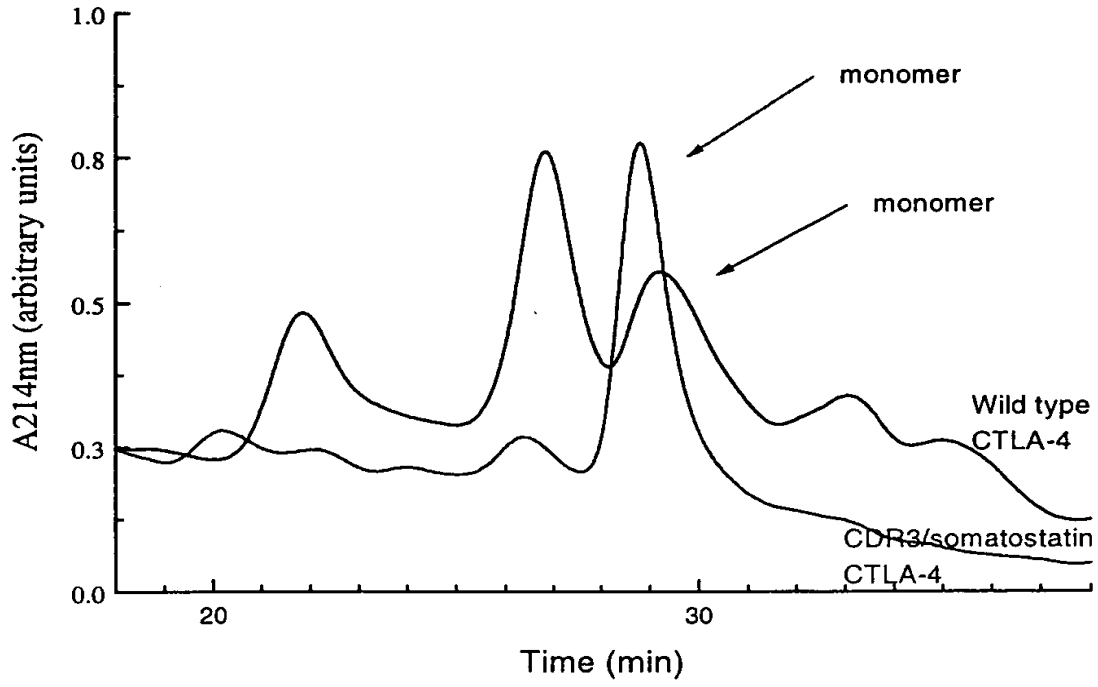
B

Epitope	*Sequence*
CDR1 Wild type	SPGK
Somatostatin peptide	AGCKNFFWKTFTSC
CDR2 Wild type	YMMGNELTF
Haemagglutinin tag	YPYDVPDYA
CDR3 Wild type	LMYPPPYYL

A. Clone designation and primary amino acid sequences of CTLA-4 STMs as expressed in plasmid vector pGC. Areas in bold represent loop regions substituted in this study. The C-terminal 21 residues of each protein consist of a dual FLAG affinity tag for protein isolation and identification.

B. Legend detailing the epitope tag sequences described in this provisional patent.

Figure 3



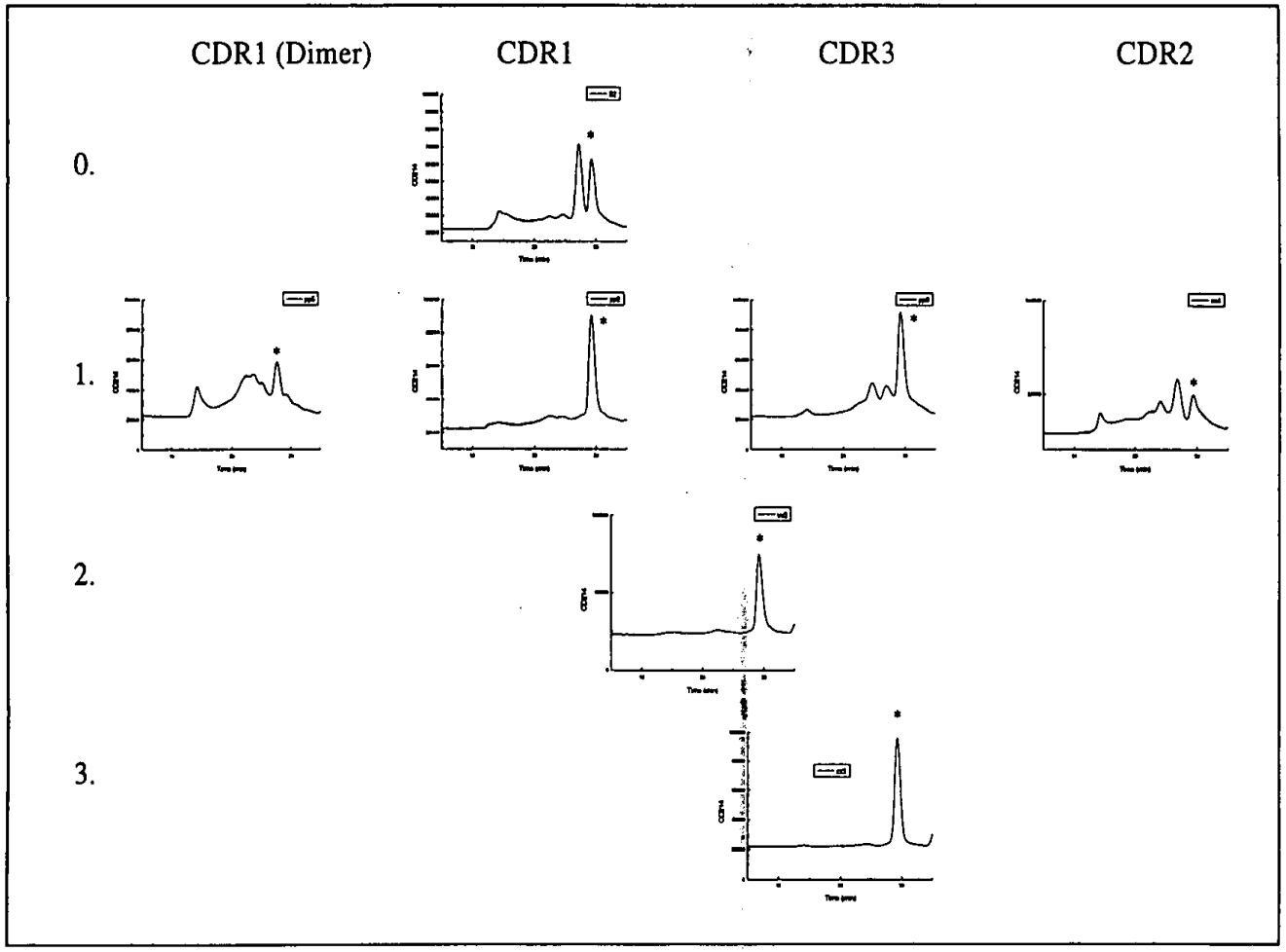
Recombinant human CTLA-4 STM proteins were expressed in *E.coli* Tg-1, purified from periplasmic extracts by affinity chromatography, and subjected to gel filtration on a superose 12 column. This graph represents a typical experiment and shows part of the elution profile of the wild type CTLA-4 STM including the following protein species: tetramer (21.86min), dimer (26.83min) and monomer (29.35min). The elution profile of STM PP8, where the CDR3 loop is replaced by the 14 residue somatostatin peptide, includes a significantly increased proportion of monomeric protein. Species present are: dimer (26.34min) and monomer (29.28min). Traces represent absorbance at 214nm and are given in arbitrary units.

Figure 4

Clone	CDR1	CDR2	CDR3	Cys^{120}	Residues	MW	Residues+Flag	MW+Flag	Elution time (min)
S2	wt.	wt.	wt.	+	115	12275	136	14638	29.35
PP2	somatostatin	wt.	wt.	-	125	13528	146	15891	29.27
PP5	somatostatin	wt.	wt.	+	135	14597	156	16960	27.63
PP8	wt.	wt.	somatostatin	-	120	12759	141	15122	29.28
VV3	somatostatin	wt.	somatostatin	+	130	14011	151	16374	29.24
XX4	wt.	HA	wt.	-	115	12272	136	14635	29.37
ZZ3	somatostatin	HA	somatostatin	+	130	14008	151	16377	29.23

Table showing designations, characteristics and descriptions of CTLA-4 STM proteins. Elution time is of monomeric proteins by gel filtration chromatography upon a calibrated superose 12 column.

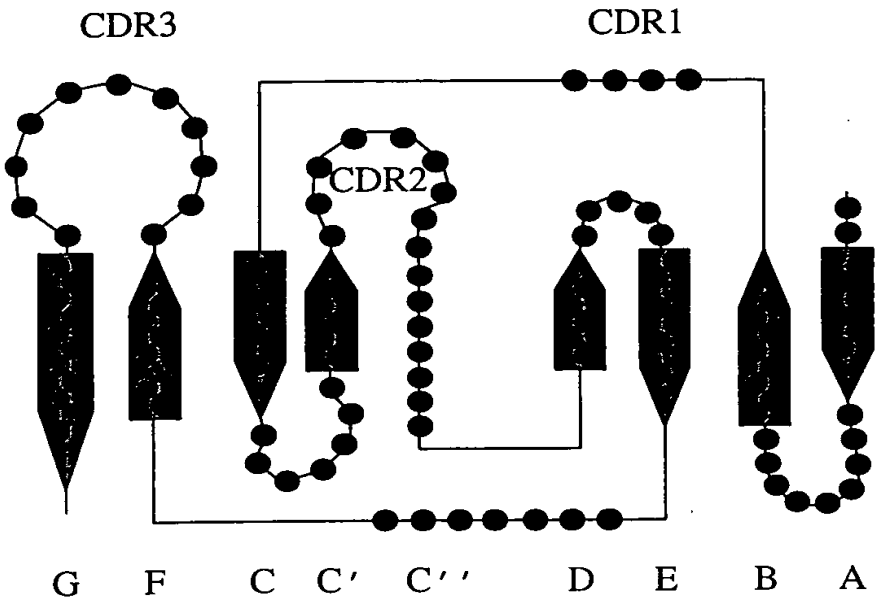
Figure 5



Results of CTLA-4 STM CDR-substitution experiments. Shown are gel filtration (superose12) profiles of affinity purified recombinant proteins, similar to drawing 3. Numbers to the left indicate numbers of substitutions; and astericeS denote monomeric peaks. STM designations are given in boxes, and correspond to those given in drawing 4.

Figure 6

A



Secondary structure of human CTLA-4 illustrating the nine B-sheet regions (numbered A-G) connected by three CDR-like loops. Refer drawing 2 for exact locations of CDR loops.

B

STM random libraries were created from the above framework by modification of CDR-like regions. Examples of such randomisations are:

Library	CDR1[1]	CDR2	CDR3
wildtype	SPGKL	YMMGNELTF	LMYPPPYYL
CDR1 library 1	XXXX[2]	YMMGNELTF	LMYPPPYYL
CDR1 library 2	XXXXXXXXCX	YMMGNELTF	LMYPPPYYL
CDR1&3 library 1	XXXXXXXX	YMMGNELTF	XXXXXXXXX
CDR1&3 somlibrary 1	AGCKNXXXXXXTSC[3]	YMMGNELTF	AGCKNXXXXXXTSC

[1] Represents the CDR loop which replaced.

[2] X represents randomisation, allowing any residue to occupy that position.

[3] Randomised loop based upon somatostatin peptide.